



2012

Do Damages Caps Reduce Medical Malpractice Insurance Premiums?: A Systematic Review of Estimates and the Methods Used to Produce Them

Kathryn Zeiler

Georgetown University Law Center, zeiler@law.georgetown.edu

Lorian E. Hardcastle

Georgetown University Law Center, leh46@law.georgetown.edu

Georgetown Public Law and Legal Theory Research Paper No. 12-186

Georgetown Business, Economics and Regulatory Law Research Paper No. 12-042

This paper can be downloaded free of charge from:


<https://scholarship.law.georgetown.edu/facpub/1130>

<http://ssrn.com/abstract=2179074>

Do Damages Caps Reduce Medical Malpractice Insurance Premiums?: A Systematic Review of Estimates and the Methods Used to Produce Them, in RESEARCH HANDBOOK ON THE ECONOMICS OF TORTS (Jennifer Arlen ed., Northampton, MA: Edward Elgar Pub 2012)

This open-access article is brought to you by the Georgetown Law Library. Posted with permission of the author.

Follow this and additional works at: <https://scholarship.law.georgetown.edu/facpub>

 Part of the [Insurance Law Commons](#), and the [Torts Commons](#)

DO DAMAGES CAPS REDUCE MEDICAL MALPRACTICE INSURANCE PREMIUMS?: A SYSTEMATIC REVIEW OF ESTIMATES AND THE METHODS USED TO PRODUCE THEM

Kathryn Zeiler^{*} and Lorian Hardcastle[‡]

Abstract

Despite common claims made in policy debates, the theoretical connection between tort reform and medical malpractice insurance premiums is ambiguous. Simple models suggest reforms such as statutory damages caps reduce premiums. More elaborate models that account for changes in physician behavior suggest caps might increase or have no impact on premiums. A number of empirical studies have been conducted to estimate the impacts of caps on premiums, and several qualitative literature reviews have attempted to draw general conclusions from the literature. No review, however, has offered a comprehensive and systematic analysis of the full set of empirical studies. This Chapter fills that gap. We provide a first glimpse at the wide methodological variations in the studies that employ regression analysis to estimate the impacts of caps on medical malpractice insurance premiums. We describe 16 empirical studies that report 197 estimates of the impact of caps on premiums. Using a theory-driven framework to develop a set of best practices, we find that little weight can be put on any one study due to broad methodological shortcomings. This Chapter highlights the need for better data and additional research on the impact of caps on premiums.

^{*} Professor of Law, Georgetown University Law Center, zeiler@law.georgetown.edu; www.georgetown.edu/faculty/kmz3/. This study is forthcoming in RESEARCH HANDBOOK ON THE ECONOMICS OF TORTS, (ed. Jennifer Arlen). Some of the work was conducted during a research fellowship at Harvard Law School's Petrie-Flom Center. Special thanks go to Jennifer Arlen, Charles R. Silver, Joshua Teitelbaum and a number of anonymous reviewers for helpful comments. Useful suggestions were provided during workshops at University of Southern California School of Law, Georgetown University Law Center, and New York University School of Law. Chinyere Ajanwachuku, Sean Dickinson, Timothy Drake, Sean Higdon and Sophia Lee provided excellent research assistance. Many thanks, as always, go to Morgan Stoddard, my exceptional law library research companion.

[‡] Visiting Scholar, University of Toronto Faculty of Law

1. Introduction

The medical malpractice liability system is designed to compensate victims of medical malpractice both for economic damages and non-economic damages, such as pain, suffering and loss of consortium. A second major function of the system is to provide incentives for health care professionals to provide reasonable care.¹ By internalizing the cost of injuries to physicians, the hope is that the threat of malpractice liability will encourage physicians to provide care that will reduce the likelihood of iatrogenic injuries.

Most providers purchase malpractice insurance to protect themselves against the risk of claims. Over the last several decades, insurance prices have fluctuated wildly, and the liability system, and large damage awards in particular, is most often pointed to as the culprit. The U.S. has faced three distinct medical malpractice insurance crises since the mid-1970s. These crises are characterized by sharp increases in premiums followed by market exit by some insurers, which reduces supply, putting additional upward pressure on prices.

Each crisis has triggered a call for major reform of the tort system. U.S. states began implementing tort reforms during the mid-1970s, with subsequent waves of reforms in the mid-1980s and early-2000s (Mello, 2006; Nelson et al., 2007). One of the most popular reforms, caps on damages available at trial, has taken center stage.² Congress continues to consider whether to implement a federal cap on non-economic damages, which all states would be required to impose on jury verdicts in medical malpractice cases (Hyman and Sage, 2011; House bill H.R. 5 (March 22, 2012)).

¹ The general standard of care is that which a reasonable physician in the same class as the treating physician would have provided. Class is determined by factors such as knowledge, skill, experience, training, and education. See e.g., *Shreeley v. Memorial Hospital*, 710 A.2d. 161 (1998).

² By Ronen Avraham's (2011) latest count, nine states impose some sort of cap on total medical malpractice damages, 26 on non-economic damages and 31 on punitive damages.

The question of whether caps reduce premiums is important. Evidence suggests that caps come with costs, such as disproportionate effects on economically disadvantaged groups and vulnerable segments of the population (e.g., Rubin and Shepherd, 2008). So if the benefits legislators promise in the form of lower premiums—and adequate physician supply³ and lower health care costs⁴—are not realized, caps result in a net loss. While legislators also argue that caps will reduce defensive medicine, recent studies indicate that doctors engage in much less defensive medicine than is commonly argued.⁵ If legislatures are left to hang their hats only on premium reductions as the primary benefit of caps, understanding the relationship between caps and premiums is critical.

While simple theoretical models predict that caps will lower payouts and thus lower premiums, more nuanced models suggest that the story might not be so simple. Models that account for the potential impact of caps on provider treatment choices and their decisions over how much effort to expend to develop and maintain medical expertise predict that caps will increase injury rates, at least under some conditions. In turn, this may cause an increase in claim rates, total payouts by medical malpractice insurers, and malpractice insurance prices.

These theories have been put to the test in a number of empirical studies. Since the early 1980s, researchers have generated empirical research studies investigating the impacts of caps on

³ See e.g., Yang et al. (2008), which summarizes the empirical literature and provides evidence that caps do not seem to impact obstetrician/gynecologists' location decisions.

⁴ Little is currently known about the impact of caps on health care costs. See Avraham and Schanzenbach (2010) for a review of the scant literature and evidence that caps reduce private health care insurance premiums by a modest amount.

⁵ Proponents of caps commonly claim that consumers benefit from the impact caps have on defensive medicine. Specifically, the claim is that caps reduce the number of unnecessary medical procedures physician prescribe solely to protect themselves from medical malpractice liability. Recent evidence, however, suggests that defensive medicine accounts for a relatively small portion of total health care costs. See e.g., Sloan and Shadle (2009), Thomas, et al. (2010). Mello et al. (2010) use a back-of-the-envelope method to estimate that defensive medicine comprised 2.4% of health care spending in 2008.

premiums. While the general consensus seems to be that caps slow the growth of premiums,⁶ a closer and more comprehensive look at the empirical literature reveals mixed results at best. A deeper look also reveals wide variation in the methods used to produce the results.

The aim of this study is two-fold. First, we examine how strongly the literature as a whole supports the claim that caps slow the growth of premiums. Using a basic “vote count” method, we find that only 61 (31%) of the 197 estimates reported in 16 studies suggest that caps reduce premiums (at the 10% significance level using one-tailed tests). While this result might seem to contradict the general conclusion that caps reduce premiums, the usefulness of reported results depends strongly on the methods employed to produce them.

Our second aim, therefore, is to gauge the level of confidence we can place on reported results. We do this by taking a detailed look at the methods used to produce results not only at the study level but also at the estimate level. While other literature reviews count studies as either supporting the caps-reduce-premiums claim or rejecting it (Mello 2010; RAND 2011), 14 of the 16 studies report numerous estimates computed using different methods. In each of the 14 studies, the results are mixed. By assessing results at the estimate level, we provide the most nuanced set of conclusions that confidentially can be drawn from the existing empirical literature.

The conclusions are not heartening. After assessing the studies along several dimensions—generalizability of results given data employed, attention to possible model assumption violations, additional sources of measurement error and controls for omitted variable bias—we find that no one study employs a consistently solid set of empirical methods. In fact, most studies come with a number of serious limitations. The bottom-line conclusion is that the

⁶ “The most recent controlled studies show that caps moderately constrain the growth of premiums.” Mello (2006) Note that Mello also summarizes studies that do not support this claim.

literature is in need of better data and more methodologically sound analysis based on comprehensive theoretical models of how various factors, including damages caps, impact premiums.

This chapter is organized as follows. Part 2 describes the existing literature, both theoretical and empirical. Part 3 provides a systematic review of the empirical literature at the estimate level. After summarizing the estimation methods used and their purposes, we assess the validity of the reported results. We find that little weight can be placed on any estimate reported in the literature given deficiencies in estimation techniques. We describe flaws common to every study and discuss the need for a better dataset to measure malpractice insurance premiums. Part 4 concludes and offers a roadmap for future research.

2. Existing Literature

A. The Theoretical Relationship Between Caps and Premiums

Theoretical accounts of the impacts of statutory damages caps on medical malpractice insurance premiums have produced varying predictions. The most basic analyses predict that damages caps will reduce recoveries and claim rates, reduce the time it takes to settle claims and reduce overall litigation expenses (Rubin, 1993; CBO, 2011). Some predict that tort reforms will create a more stable liability environment with less underwriting uncertainty (Born and Viscusi, 1998; Barker, 1992).⁷ These effects, it is argued, will lead to a reduction in medical malpractice insurance premiums.

Alternative accounts of the impacts of caps on premiums consider second-order effects and whether caps would, in fact, bind recoveries. For example, Hyman et al. (2007) use data on

⁷ Note that these studies acknowledge that the causality may be reversed—instead of reforms improving insurer profitability, insurer behavior may determine a states' willingness to enact reforms. The mechanisms through which tort reforms influence premiums are important to understand so that empirical models can account for the complexities.

Texas closed medical malpractice claims to explore why amounts plaintiffs collect are so often well below the amounts juries award. Their results suggest that the discrepancy is due in part to the fact that plaintiffs rarely collect the portion of damage awards that exceeds the defendant physicians' per occurrence policy limit. The authors argue that limits on collectability of amounts of the policy limit can "substantially mute the real-world impact of statutory caps" (Hyman et al., 2007, p. 53). Others have suggested that plaintiffs' attorneys react to damages caps by arguing more vigorously for damages that are not capped, often economic damages such as lost wages, and that juries simply substitute higher damages of the uncapped varieties for capped damages (Sharkey, 2005).

Others have considered the impacts of caps on incentives for providers to engage in costly activities that reduce the likelihood of patient injuries. Zeiler (2003) argues that although caps might reduce average payouts, they also might increase the number of claims, thereby failing to reduce overall costs of covering physician liability. As caps reduce exposure to liability, the argument goes, the benefits physicians (and the managed care organizations that influence their treatment choices) enjoy from providing expensive treatments are diminished. At the margin, this theory predicts that caps lead to more iatrogenic injuries. Similarly, Arlen and MacLeod (2005) predict that caps will reduce physician investment in costly efforts to maintain and develop expertise. Thus, total medical malpractice insurer losses will depend on the influence of caps on claim rates relative to average claim payouts. While total losses might decrease, these theories predict that, under some conditions, they will remain stable or increase.⁸

B. Empirical Tests of the Predicted Impacts of Caps on Premiums

The temporal and geographic variation in tort reform implementation has allowed researchers to compare premiums pre-cap and post-cap in states that implemented caps and to

⁸ Rubin and Shepherd (2008) find that caps increase female death rates.

compare states with caps to those without caps. Researchers have employed the information to estimate the impact of reforms, including damages caps, on malpractice insurance premiums. As this evidence accumulated, researchers began to produce literature reviews, drawing general conclusions from the numerous studies. Our search⁹ revealed seven relevant reviews: OTA, 1993; CBO, 2004; Kessler, 2006; Kane and Emmons, 2005 and 2007 (discussed as a single study); Nelson et al., 2007; Studdert and Mello, 2004, Mello, 2006, Mello and Kachalia, 2010, and Kachalia and Mello, 2011 (discussed as a single study¹⁰); and RAND, 2011.¹¹

Appendix 1 summarizes the scope of each literature review. The reviews include different sets of primary studies, although none covers all 16 studies included in our analysis. Some describe the method used to include particular studies and exclude others, most of which involve an evaluation of the strength of the empirical methods the primary study authors used and whether and by whom the study was published. Most reviews draw general conclusions about the impacts of several tort reforms, while some focus on the impacts of caps. Some reviews go beyond our variable of interest—premiums—to consider the impacts of tort reform on payment size, claim rates, defensive medicine, physician supply, quality of care, defense costs, and health expenditures. We focus solely on the association between caps and premiums, the question most widely studied and of central interest in current political debates.

⁹ Our goal was to identify all literature reviews drawing general conclusions about the impacts of damages caps on premiums from studies that employed regression analyses. We searched several databases (PubMed, EconLit, SSRN, ProQuest Theses and Dissertations, Catalog of U.S. Government Publications, Academic Search Primer, Google Scholar, ABI/Inform Global, Google Books, Westlaw and Index to Legal Periodicals) generally using the following search term: (“tort reform” OR “damage cap” OR “damage caps” OR “damages caps”) AND (“literature review” OR “systematic review”). Because literature reviews (particularly government reports) tend not to self-identify as “literature reviews,” we searched for papers citing the primary tort reform studies. Finally, we searched the bibliographies of the literature reviews to find references to other reviews.

¹⁰ Studdert and Mello, 2004 is a shorter, less detailed version of Mello, 2006. Mello and Kachalia, 2010 includes more recent information. Kachalia and Mello, 2011 is a shorter version of the 2010 paper, with little additional information.

¹¹ This literature review used Mello’s 2006 methodology but addressed more recent cap studies. There is significant overlap between the RAND review and Mello’s 2010 update of her 2006 paper.

The two most recent and comprehensive reviews to date conclude that damages caps likely reduce premiums. The first by Mello and co-authors (2006 and 2010) is the most sophisticated review. The 2010 version includes eight primary studies. The authors set out the following criteria for assessing the rigorousness of a study: a sufficiently comprehensive data source, low potential for error or bias, appropriateness of methodology, adequacy of controls for confounding variables, adequacy of sample size, and appropriateness of interpretation and conclusions. The authors conclude that more recent data suggest that caps reduce premiums by 6-13%, although the authors cite numerous studies that report no effect on premiums.

The second recent review by RAND (2010) piggybacks off the work of Mello and her co-authors. This brief review updates Mello's 2006 findings, employs a similar methodology for study inclusion, and corroborates Mello's conclusions. Summarizing her earlier findings, the RAND review concluded that four studies (two of which were methodologically strong) showed a link between caps and reduced premiums and four studies (one strong) did not.

While both reviews are helpful in demonstrating the variation in reported results, they come with an important limitation. Specifically, both reviews classify studies "as having significant findings if any specification of a damages cap variable was statistically significant." (Mello, 2006, p. 10) In doing so, they brush aside ubiquitous within-study variation in results.¹² None of the primary studies included in the Mello et al. and RAND reviews reported consistent results; at least half of each study's results failed to support the claim that caps reduce premiums.¹³

¹² Mello's review lists Zuckerman et al (1990), Danzon et al. (2004), Thorpe (2004) and Viscusi & Born (2005) as support for the claim that caps reduce premiums even though 3 of 6, 4 of 6, 4 of 6 and 9 of 12 results, respectively, failed to support the claim at the 10% significance level (see Table 1 below). Despite these mixed results, Mello concludes (from Danzon et al, Thorpe and Viscusi et al.) that "studies based on data from the 1990s and the early years of the current malpractice crisis *consistently* found that caps had a modest but statistically significant constraining effect on premiums during this period..." (p. 12; emphasis added)

¹³ Section 3, *infra*, details the individual results reported by each study.

Our goal is two-fold. First, we aim to present a more comprehensive and detailed summary of results from primary research studies, taking into account the widespread within-study variation. Second, we attempt a more nuanced evaluation of the methodological soundness of studies based on the use of theory-driven best methods.

3. Systematic Literature Review

A. Overview of Primary Study Results

Our first step involved collecting all relevant empirical studies. As described in Appendix 3, we tailored our search terms to particular research databases. We reviewed each relevant result for additional studies our searches had not captured and for mentions of unreported regressions (e.g., robustness checks). We were unable to obtain results from unreported regressions.¹⁴ Our analysis includes all of the studies from the literature reviews and additional studies not included in any reviews (in bold type in Table 1). We identified 16 studies reporting 197 estimates of the impact of caps on premiums.

We begin with the most basic inquiry: considering all available evidence from regression analyses that estimate the association between caps and premiums, how strongly does the evidence support the common claim that caps reduce premiums? Summary statistics computed using the entire set of results suggest a closer look is warranted. Only 61 (31%) of the 197 reported results support the claim that caps are associated with a statistically significant decrease in premiums (10% significance level for one-tailed test). Forty-nine of the 197 results (25%) support the claim at the 5% significance level.

¹⁴ Sloan, 1985; Danzon et al., 2004; Viscusi and Born, 1995; Viscusi and Born, 2005: no longer have records. Barker, 1992: would take considerable time to review methodology and find unreported regression records. Crain et al., 2009: no access to records. Grace and Leverty, 2011: paper in progress so other results not yet available. Zuckerman et al., 1990; Thorpe, 2004: no response.

Researchers have used three different proxies for premiums. Some studies use *premiums paid by physicians* of certain specialties for policies with identical policy limits and duration. For example, Sloan (1985) uses survey data on premiums paid by general practitioners, ophthalmologists, and orthopedic surgeons for policies with limits of \$100,000 per occurrence and \$300,000 over the coverage period (“\$100,000/\$300,000 policies”). Bhat (2001) uses premiums charged by a single multi-state medical malpractice insurer for physicians of the same risk class and for \$1M/\$3M policies. Others use *aggregate premiums* written (Blackmon et al., 1991; Viscusi et al. 1993) or aggregate premiums earned (Viscusi and Born, 1995; Born and Viscusi 1998) as a proxy for price.¹⁵ Finally, several authors employ *loss ratios*—the ratio of losses incurred by an insurer to total premiums earned by that insurer—as a proxy for price. The inverse of this ratio represents the effective price of insurance: the amount physicians pay for each dollar the insurer indemnifies.¹⁶ Thus, if caps are correlated with increased loss ratios, we can conclude that caps reduce the effective price of insurance.

Similarly, primary study authors have studied a variety of damages caps. Some study the impact of caps on non-economic compensatory damages (e.g., pain and suffering). Others study the impacts of punitive damages caps, while others investigate the impact of caps on total damages, or they group together caps of various sorts.¹⁷ Some have posited different impacts for different types of caps (e.g., Sharkey (2005) argues that punitive damages will have a smaller

¹⁵ “Premiums written” is the total premiums generated from all policies written by an insurer within a given period of time (cash basis of accounting for premiums). “Premiums earned” is the portion of the total premium amount corresponding to the coverage provided during a given time period (accrual basis of accounting for premiums).

¹⁶ The loss ratio “can be used as a measure of the relative price of insurance because it reflects the proportion of premium dollars that are returned in the form of loss payments to those insured.” (Barker, 1992, p. 145); “[O]ne can view the loss ratio as a measure of the inverse of the ex post unit price of insurance.” (Viscusi and Born, 2005, p. 38)

¹⁷ This group includes cap of any sort, cap on total damages or immunities from liability, caps on the total amount physicians pay but not the amount patients received, caps on economic and non-economic damages, caps on economic and/or non-economic damages, and caps on economic or non-economic damages.

impact on premiums relative to non-economic and economic caps because punitive damages are rarely awarded in medical malpractice cases).

Table 1 provides a breakdown of the result counts by premium proxy type. It also lists the publication type of each study and the years covered by the premiums data and the caps data. The breakdown reveals a substantial difference in the pattern of results depending on the proxy used for premiums. When premiums per physician or aggregate premiums are employed, 55-57% of the results are not significant at the 10% level, 7-10% of results are significant at the 10% level (but not the 5% level) and 34-35% are significant at the 5% level. When researchers use loss ratios as a proxy for premiums, none of the results supports the claim that caps reduce premiums at the 10% level.

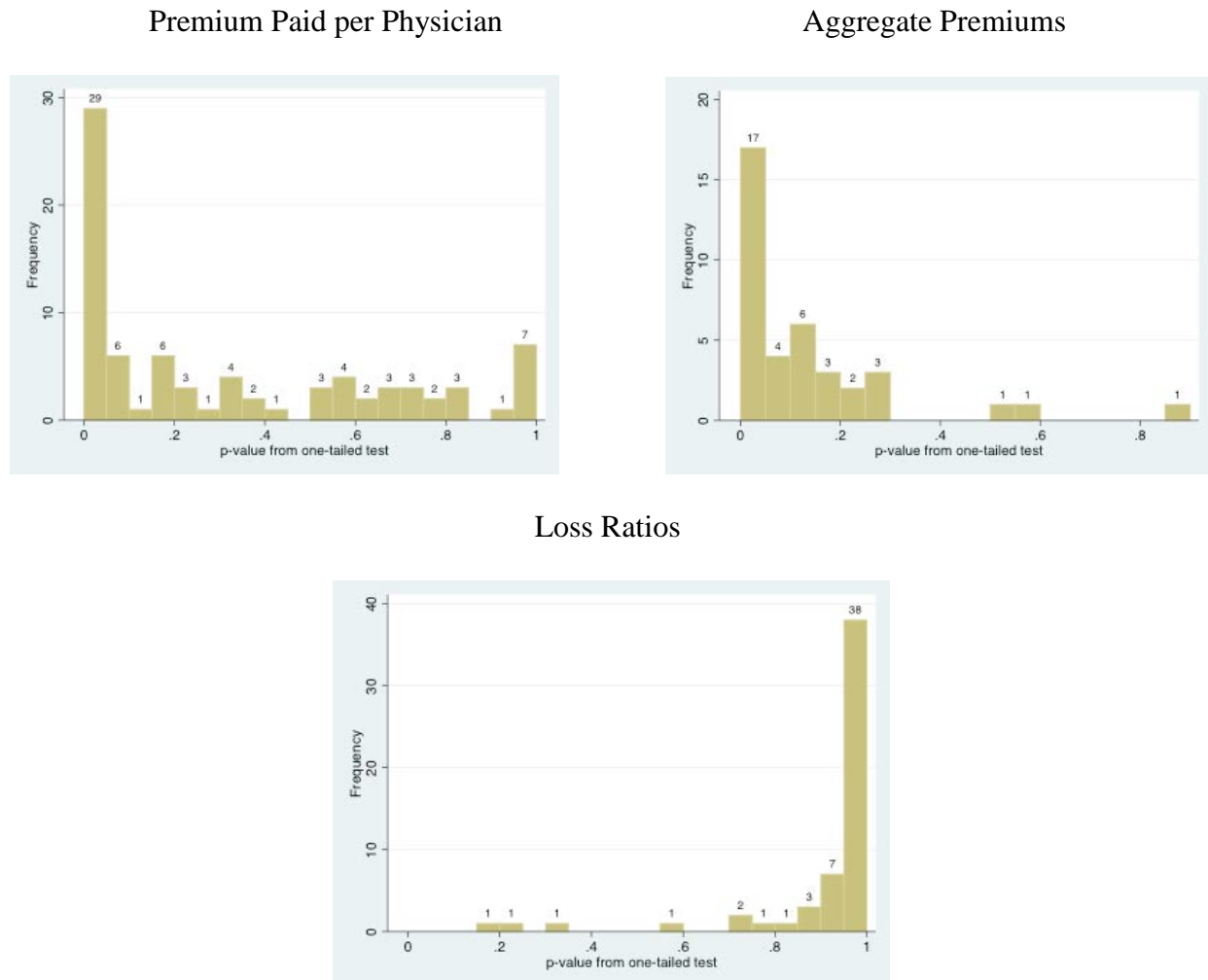
Table 1: Features of primary studies

Study (first author)	Publication type	Premiums period	Caps period	# of esti- mates	Premium Paid/Doc			Aggregate Premiums			Loss Ratio		
					Not s.s.	5-10%	5%	Not s.s.	5-10%	5%	Not s.s.	5- 10%	5%
Sloan 85	PRJ	1974-1978	1974-1978	14	13		1						
Zuckerman 90	PRJ	1974-1986	1974-1986	6	3		3						
Blackmon 91	book chapter	1985, 1988	1986	4				2		2			
Barker 92	PRJ	1977-1986	1975	4							4		
Viscusi 93	PRJ	1988	1985-1987	6				3			3		
Viscusi 95	PRJ	1985-1991	1985-1987	6				3			3		
Born 98	Brookings Inst.	1985-1991	1985-1987	42				6	3	12	21		
Guis 98	PRJ	1976-1990	1976-1990	1				1					
Bhat 01	book	1990-1994	1990-1994	4	2		2						
Zeiler 03	dissertation	1991-2001	1991-2001	4	1	1	2						
Danzon 04	Brookings Inst.	1994-2003	1994-2003	6	4		2						
Thorpe 04	PRJ	1985-2001	1985-2001	6	1		1	1		1	2		
Viscusi 05	PRJ	1984-1991	1984-1991	12				3	1	2	6		
Kilgore 06	PRJ	1991-2004	1991-2004	45	23	5	17						
Crain 09	PR Inst	2004, 2006	2004, 2006	1			1						
Grace 11	unpublished	1985-2005	1985-2005	36				13	2	3	18		
Total				197	47 (57%)	6 (7%)	29 (35%)	32 (55%)	6 (10%)	20 (34%)	57 (100%)	0 (0%)	0 (0%)

Notes: PRJ = peer-reviewed journal; PR Inst = Pacific Research Institute; Not s.s. = not statistically significant at the 10% level (one-tailed test); 5-10% = significant at the 10% level, but not at 5%; 5% = significant at the 5% level. Bolded studies were not included in any prior literature review.

The average p-value produced by 194 one-tailed tests of significance is 0.47 (standard deviation = 0.40).¹⁸ Figure 1 displays the distribution of observed or estimated one-tailed p-values for results produced using different proxies for premiums.

Figure 1: Distribution of one-tailed p-values by premium proxy



The Figure demonstrates that conclusions drawn from the empirical literature depend greatly on the proxy. Although the majority of results produced using per-physician premiums and

¹⁸ Thorpe (2004) did not provide sufficient information to estimate p-values related to three of the six results he reported.

aggregate premiums are statistically insignificant at the 10% level, these proxies are much more likely to produce statistically significant results than loss ratios. The results produced using loss ratios support the claim that insurers collect *more* in premiums per indemnified dollar after caps are implemented.

As Mello et al. (2006, 2010) point out, vote counts can be misleading if some of the studies producing results are not methodologically sound. In the following section we describe the methods commonly used by primary study authors to obtain valid results. That section is followed by a discussion of what we can infer from the literature given the methods employed.

B. Estimation Methods

Had policy makers implemented caps randomly over states and years while holding all else constant, estimating the impact of caps on premiums merely would entail comparing the difference in premiums before and after the implementation of caps in the treatment states relative to differences in the control states over the same time period. Estimation is not so simple given that states did not implement caps randomly from year to year and all else was not held constant. Therefore, the validity of results depends on the use of methods to control for measurement error and whether the necessary assumptions of the chosen empirical model are satisfied. In addition, the data employed will impact generalizability of reported results. This subsection describes methods necessary to produce valid and generalizable results.¹⁹

1. Model and Sample Selection

The most preliminary choices researchers make involve selection of the empirical model, the sample size, the level of observation (e.g., state or insurance company), and the time period covered by the data.

¹⁹ Note that this section is not intended to be a comprehensive list of such necessary methods. The discussion is tailored to the methods employed in the 16 original studies. Our goal is to provide a sense of some of the common methods for properly modeling the impacts of policy changes on outcomes.

Every original study uses one of three basic model types: ordinary least squares (OLS), weighted least squares (WLS) and quantile regression (QR). OLS models estimate linear relationships by minimizing the sum of squared errors (e.g., differences between observed premiums and values predicted by the linear approximation). WLS models employ OLS techniques but weight the observations to take into account the possibility that not all observations are equal (e.g., Barker (1992) assumes the variance of error decreases proportionally as risk exposure increases and, thus, weights observations by a function of the premium volume by state). QR models are used to account for the possibility that caps differentially impact firms (e.g., firms of various sizes or profitability levels (Born et al., 1998)) and that outliers might impact estimates.

Studies also vary with respect to the **level of observation**. Some researchers use panel data with observations at the state-year level (Gius, 2004); others use panels with observations at the insurer-state-year level (Zuckerman, 1990; Kilgore, 2006); and some employ cross-sectional datasets, with all states observed in a single year (Viscusi and Born, 2005).²⁰ Panel datasets are known to produce more precise estimates, partly because they come with larger **sample sizes**, which allow for higher **degrees of freedom**.²¹ Generally, the power of the statistical test increases as the number of degrees of freedom increases. The higher the power, the higher the likelihood of detecting an effect if one exists.

Along the same lines, the **number of years** covered by the data will impact the precision of the results. The primary studies cover a range of time periods for which caps are coded and for which insurance outcomes are recorded. For example, Viscusi and Born (1995) and Born et al. (1998) considered whether variation in caps enacted during the years 1985-1987 explains any of

²⁰ All studies that cover multiple years adjust premiums for inflation.

²¹ The number of degrees of freedom is equal to the sample size minus the number of independent variables included in the regression equation minus 1 (for the constant term).

the variation in premiums across insurers for the years 1985-1991. Guis (1998) used a panel dataset, which included variation in both caps and premiums for the years 1976-1990. Earlier studies typically consider a substantially smaller range of years (e.g., Zuckerman (1990) argued that Sloan's (1985) results might be impacted by the fact that his data cover only four years).

2. Generalizability

The extent to which we can generalize from reported results depends on the characteristics of the data employed and whether those characteristics generalize to the population about which claims are made. All 16 original studies draw general conclusions about the impact of damages caps on insurance premiums. Given the nature of the data, however, such claims might be too broad in some cases.

First, reported results vary over the **time periods** studied. This variation might be due to changes over time in impacts caps have on prices. For example, caps that are not inflation-adjusted might become more binding over time, leading to a decrease in prices). Alternatively, it might result from unobservable changes to markets or the liability system that are unrelated to but correlated with caps. For example, changes in the average medical malpractice policy size purchased, which is typically unobservable (i.e., the data are difficult to obtain), might impact premiums.²² If unobservable features change over time, putting the greatest weight on recent studies might be most prudent (assuming the current environment is most similar to the environment during the most recently studied years) (Mello, 2006).

Second, and more importantly, the **data employed to measure premiums** should allow us to draw general conclusions from the results. To start, the data should act as a good proxy for premiums. If we have reason to believe that caps impact premiums differently than the variable

²² Zeiler et al. (2007) report that Texas physicians that have made payments to claimants have purchased policies with lower policy limits over time.

used as a proxy, generalizing from the proxy to premiums is problematic. Alternatively, the proxy might be a good one but might represent a small slice of the population of interest. Again, if we have reason to believe that caps might impact premiums of different types of policies differently, then we should be careful about generalizing from a subset of policy types to all policy types. The data used in the 16 original studies come from a variety of **sources**, cover a variety of **provider types** and include a variety of coverages (i.e., **policy limits**), which are sometimes **controlled** in the analysis. In some cases inferences about the impacts of caps on premiums are restricted to policies with a certain level of coverage (e.g., \$100,000 per occurrence and \$300,000 annual aggregate). In other cases, policy limits are controlled so that variation in premiums related to variation in coverage is separated from premiums variation caused by variation in caps. In the latter case, the estimated impact of caps on premiums is interpreted as the average impact across all policy types.

Finally, some studies estimating the impacts of caps on non-economic damages identify separately the impacts of **caps of different types**. Some studies consider the impacts of caps of different sizes (e.g., non-economic caps cannot exceed \$250,000). Others consider the impacts of permanent caps separately from the impacts of temporary caps, which eventually are overturned or repealed (Grace and Leverty, 2011).

3. Model assumption violations²³

The primary studies employ a wide variety of empirical modeling techniques. As mentioned, OLS is the most commonly employed regression model. This model requires several assumptions to be satisfied. The most widely violated assumptions include non-normal distribution of error terms, heteroskedastic error terms, serial correlation of error terms, and

²³ The technical details included in this section are explained in basic econometrics textbooks. E.g., see Wooldridge (2008) and Greene (2002).

simultaneity.²⁴ For our purposes, the technical details of these assumption violations are unimportant. We focus instead on whether primary study authors properly attended to them.

Non-normal distribution of error terms. The basic linear regression model assumes that error terms are normally distributed. A non-normally distributed dependent variable can result in non-normally distributed error terms. When a dependent variable takes on only positive values (e.g., premiums), a common approach used to nudge the error terms towards a normal distribution is to log-transform the dependent variable. While this transformation might result in error terms that satisfy the normality assumption, it is not a guaranteed fix. Alternative transformations might solve the problem.

Heteroskedastic error terms. If errors term variances are not constant across different values of the independent variables, our confidence in hypothesis test results is compromised. Methods used to address this problem include computing heteroskedastic-consistent standard errors (robust standard errors more generally) or computing standard errors using a bootstrap resampling technique. Log-transforming the dependent variable gets around the problem in some cases but is not a guaranteed fix. WLS regression is sometimes used to control for heteroskedasticity, but the weights must be chosen carefully and with heteroskedasticity specifically in mind.²⁵

Serial correlation of error terms. Most primary studies use panel datasets, which are comprised of data collected from each state (or several companies within a state) over a number

²⁴ The basic linear regression model also assumes that the relationship between the dependent variable and every continuous independent variable is linear. Since most independent variables employed in the original studies are binary, this assumption is not one that requires much attention. It should be noted, however, that no study mentions verifying linearity when continuous variables (e.g., income, insurance market size) are employed as controls. If the model does not properly capture the relationship between premiums and the controls, the estimate of the impacts of caps on premiums might be biased.

²⁵ Only Barker (1992) uses weighted least squares regression specifically to address heteroskedasticity. She relies, however, on the untested assumption that the variance of the error decreases proportionally as the number of risk exposures increases.

of years. This type of data often violates the assumption of uncorrelated error terms. While serial correlation does not lead to biased estimates, it can lead to faulty conclusions about statistical significance. Tests have been designed to detect the problem, and, if the problem presents, standard errors must be corrected (e.g., Bertrand, et al. (2004)) or a model other than the standard OLS should be used (e.g., auto-regressive model, moving average model, feasible generalized least squares) (King and Roberts, 2012). As with every model, the alternative models will produce valid results only if the model's assumptions are satisfied. In addition, if the problem does not exist and a correction is adopted, the results could be invalid.

A second common method for addressing possible serial correlation is by clustering standard errors by group. Dependence among observations often arises in data with a group structure. For example, premiums set by the same company in the same state might be influenced by similar factors over time, so that relative high premiums in one year might imply relatively high premiums in other years. This might not be a perfect solution, however, if the number of clusters is relatively small or the number of years covered by the dataset is small. Typically if standard errors are clustered by state and all 50 states are included in the dataset, results are likely to be valid. (Angrist and Pischke (2009), p. 323)

Simultaneity. The basic linear regression model also assumes that independent variables are exogenous, which, in the present context, requires that changes in premiums over time do not influence the likelihood that caps are implemented. The probability that caps are implemented, however, likely depends on trends in premiums; changes in premiums and the likelihood of cap implementation theoretically impact one another simultaneously. If the simultaneity assumption is not satisfied, the researcher can address it by using a number of possible methods including an

instrumental variables approach, propensity score matching (Rosenbaum and Rubin, 1983) and randomization inference (Donohue and Ho, 2007; Imbens and Rosenbaum, 2005).

4. Additional sources of measurement error

Obtaining correct estimates of the impacts of caps on premiums requires a fine-tuned coding method. Zuckerman (1990) notes that better estimates are produced when tort reforms are coded using **effective dates** rather than enacted dates, as these dates can differ in some cases. He also points out the importance of coding tort reforms in a way that **accounts for legislative repeals and court decisions** to strike down reforms as unconstitutional. Similarly, Grace and Leverty (2011) hypothesized that the presence or absence of caps is insufficient to predict insurer behavior—one must also consider the **likelihood of judicial nullification**. They consider temporary reforms separately from permanent reforms.

In addition, researchers sometimes account for the possibility of insurers' **delayed response** to legal reforms due, for example, to uncertainty around possible effects of the change (Danzon, 2004; Zuckerman et al., 1990).²⁶ Several authors suggest that no-effect results might be caused by lack of attention to the possibility that reforms often take several years to have an impact given the long-tailed nature of the legal claims.²⁷ Including **lags** of tort reform variables in the regression model²⁸ is a common method of accounting for a possible delayed response.

Others have suggested the opposite reaction—that insurers actually **anticipate changes** to tort law and adjust prices prior to the change. If actors anticipate changes to tort law in some future year, premiums in the current year might reflect these predictions (Malani and Reif, 2011).

²⁶ As Born and Viscusi (1998) hypothesize, insurer losses will first reflect the impact of tort reforms, but premium changes might lag for several reasons including the need to obtain regulatory approval for rate changes.

²⁷ For example, Thorpe (2004) cites unpublished data from one large insurer revealing that nearly 12% of claims took at least eight years to resolve.

²⁸ Rather than, or in addition to, including an independent variable for tort law in effect in the year premiums are measured, an independent variable for tort law in place in the previous year(s) is included to capture the impacts of reforms passed in previous years.

Grace and Leverty (2011) theorize that while anticipation of the implementation of a damages cap will not impact prices prior to implementation, insurers might adjust insurance prices in anticipation of judicial strike downs of caps. The same might be posited for legislative changes to previously implemented caps. Grace and Leverty account for this possibility by separately estimating the impacts of permanent caps—those that are not deemed unconstitutional by the court—and temporary caps—those eventually overturned by the court. Others suggest incorporating “leads” of tort reform variables,²⁹ although knowing whether the evidence supports anticipation or simultaneity is difficult.

Researchers have also recognized the importance of properly **weighting observations** to obtain accurate average effect sizes that reflect, for example, market share (e.g., when insurer-level premiums data is used to construct the dependent variable) and market size (e.g., when state-level premiums data are used). Weighting ensures that, when computing estimates of the *average* impacts of caps on premiums, more weight is placed, for example, on firms with larger market shares or on states with larger insurance markets. While failing to properly weight observations does not impact the statistical significance or sign of the result, it produces a biased estimate of the average impacts of caps.

Finally, OLS estimates are known to be sensitive to **outliers**, sometimes called influential observations. An observation is influential if the estimate changes by a practically large amount when the observation is dropped from the dataset. All else constant, the likelihood of the presence of outliers in the dataset is reduced as the sample size increases. Log-transforming the dependent variable can make the estimate less sensitive to outliers, but influential outliers might

²⁹ Rather than, or in addition to, including an independent variable for tort law in effect in the year premiums are measured, an independent variable for tort law in place in the future year(s) is included.

remain. Similarly, estimates produced using QR models are less sensitive to outliers, but these models do not guarantee a remedy.

5. Controlling omitted variable bias

In a perfect world, we would estimate the impact of caps on premiums by randomly assigning states to treatment (caps) and control (no caps) groups and measuring the difference, post-treatment, between average premiums across the two groups. The function of randomized assignment is to generate two groups of states with roughly the same distributions of characteristics so differences in premiums after caps are imposed can be attributed to the caps and not to differences in other covariates that might be correlated with the imposition of caps. Controlling for these covariates reduces the possibility that a portion of the impact the model attributes to caps should actually be attributed to other characteristics of the treatment states.

Given that caps were not randomly assigned to states, researchers use methods to control both observable and unobservable covariates. Several features of insurance markets and the liability system other than caps impact premiums. If one or more of these features is correlated with the implementation of caps and is left out of the analysis, the estimated impact of cap on premiums will be inaccurate. Methodologically sound studies employ a number of controls to separate the effects of variables other than damages caps on premiums to reduce omitted variable bias.

A great deal of variation across the 16 empirical studies arises from different choices over control variables. All studies control for other tort reforms (e.g., **collateral source rules, modifications to joint and several liability, patient compensation funds, periodic payment reforms, time limits on claims including statutes of limitations and repose, implementation of pretrial screening panels, arbitration rules, ad damnum reforms, expert witness reforms,**

and prohibition against punitive damages insurance), but they include different sets of reforms.³⁰ These controls are especially important as reforms tend to be passed as packages, which increases the likelihood of correlation between caps and other types of reforms and, in turn, the likelihood of omitted variable bias absent controls.

Researchers often mitigate omitted variable bias by including controls for features other than tort reforms that are either known or suspected to impact premiums and that might be correlated with the passage of damages caps. These include controls for potential claimant demographics (e.g., **urbanization level, income per capita, citizen ideology**), **exposure** to health care services (e.g., **number of surgeries per capita, number of physicians per capita**), insurance market characteristics (e.g., **market concentration, insurer organizational form, size of insurance market, intensity of insurance price regulations, whether the state runs a joint underwriting association, rates of return on investments**), and attorney market characteristics (e.g., **restrictions on legal fees, number of attorneys per capita**). The magnitude of impacts on results depends on the strength of correlations between damages caps and each control variable.

Some researchers employ techniques to mitigate omitted variable bias caused by unobservable variables. For example, Guis (1998) argues that results from previous studies are biased because they do not control for unobservable differences across states and years. **Cross sectional fixed effects** are used to control variables that are constant across time but vary across units (states or insurers). **Time fixed effects** control for variables that are constant across units but vary over time. Some include a **linear time trend** to isolate general trends in premiums over

³⁰ In some cases, researchers (e.g., Viscusi (1993)) coded binary variables to track whether states enacted any of a number of different tort reforms (e.g., any number of **pro-defendant reforms, any reforms in place prior to premiums period, any reforms in place other than caps**). This method produces more noise than methods that include separate controls for covariates.

time that might confound estimates of the impact of caps on premiums. In some cases, researchers include a **lagged dependent variable** for the same reason, although this technique is problematic when a panel dataset is used. (Hsiao, 2003)

Finally, a common technique used to estimate the impacts of policy changes is **difference-in-differences** (DID) models. The method is considered “quasi-experimental”—jurisdictions that implement the policy comprise the treatment group and non-implementers act as the control group. The function of the control group is to difference out any impact on the dependent variable caused by contemporaneous changes that occurred in all jurisdictions at the time of the policy change. The standard DID model incorporates cross sectional (e.g., state) fixed effects and year fixed effects along with a set of observed covariates as additional controls.

Importantly, validity of DID model estimates depends heavily on two assumptions that often are overlooked. First, the model assumes that the policy was randomly assigned across state-years. This assumption is violated if features of the environment not accounted for in the model impact the likelihood that a state implements the policy³¹ or if the policy is likely to remain in effective in an adopting state after the first year of implementation (Bertrand et al., 2004).³² Second, the model assumes that policy implementation does not impact the covariates. If, for example, caps impact the number of physicians per capita, thereby increasing physician supply and decreasing the population by affecting mortality, the estimates of the impacts of caps on premiums will suffer from post-treatment bias (Ho, 2005).³³

³¹ E.g., if premiums naturally cycle over time and caps are adopted only when premiums reach a certain threshold, then changes in premiums might be attributed to caps when they are actually a function of the underwriting cycle. (Donohue and Ho, 2007).

³² DID models assume that the likelihood of a reform being in place in one state-year is independent of whether the reform is in place in any other state-year. In contrast, a reform that is in place in one state-year is likely to be in place the following year in that state.

³³ Including and excluding possibly tainted covariates to assess sensitivity might not help (Rosenbaum, 1984).

With these guidelines in mind, we now assess the validity of the estimates reported in the literature.

C. Validity Assessment

Tables 5-7 summarize the methods employed to produce each estimate reported in the literature.³⁴ Tables 5a-c include all results from analyses that group together various types of damages caps (e.g., caps on non-economic damages, punitive damages, total damages, etc.). Table 5a includes estimates of the impact on premiums paid per doctor. Table 5b includes estimates on aggregate premiums, and Table 5c on loss ratios. Each row in the table represents a single estimate from a single study. We report the estimate's p-value³⁵ from a one-tailed test of significance with an alternative hypothesis that the coefficient is less than zero (i.e., caps reduce premium growth), the sample size, degrees of freedom, and the methods used to produce the result. Tables 6a-c and 7 are similarly constructed but include estimates of the impact of non-economic damages caps and punitive caps, respectively. Although many lessons can be drawn from analyzing the results at the coefficient level through the lens of methodological soundness, we focus on four general take-aways.

1. Premiums data severely limits generalizability of results

³⁴ We employed a multi-step process to compile information about the primary studies. Research assistants trained by us completed the initial coding. We reviewed their coding, making corrections as necessary. We also added variables that were not identified during the initial coding process. A research assistant then checked every element of the revised dataset against the primary studies, noting any inconsistencies. We reviewed flagged items and made the necessary corrections. Appendix 4 contains a list of relevant variables.

³⁵ The p-value allows us to draw inferences about the strength of the evidence for or against the tested hypothesis. We consider p-values from one-tailed tests, which reveal the likelihood of observing a result as extreme as the observed result assuming caps do not decrease premiums. The lower the p-value, the more confidence we have in concluding that caps reduce premium growth. The general convention is to use a 0.05 cutoff. A thick line separates results using this convention.

Where the primary study did not report p-values, we computed them using coefficients and standard errors. In the few cases in which only significance levels were reported, we used the average of all possible p-values assuming a uniform distribution (0.55 was used for results insignificant at the 10% level ($0.1 + (1-0.1)/2$), 0.075 for results significant at the 10% level ($0.05 + (.1-0.05)/2$), etc.). For three of the 197 results (Thorpe, 2004), we could not estimate p-values due to insufficient information.

The **nature of the premiums data** researchers employed gives rise to important and yet underappreciated limitations. None of the datasets employed in any of the primary studies allows generalized claims about the impacts of caps. Resulting complications depend on the data employed.

Changes in aggregate annual premiums over time might be due to changes in prices, and/or changes in the number of policies sold, and/or changes in the types of policies sold (e.g., policies with varying coverage limits and deductibles). Thus, in studies that use aggregate annual premiums written or earned by state or by company, separating changes in price from changes in quantity sold from changes in the mix of policy types sold is impossible.³⁶ Aggregate premium measures tend to be better proxies for gross revenue and business growth.

³⁶ “Total premiums reflect not only the price of insurance but also the quantity of policies sold. We cannot tell if an increase in total premiums means that doctors are paying more for their insurance, or if the insurer is just selling more policies.” Mello, 2006, p. 23.

Table 5a
Any cap; Premiums paid by doctor

[illegible]

KEY: **OLS** = ordinary least squares; **wls** = weighted least squares regression; **quant**: 10 = most profitable or largest firms, 90 = least profitable or largest firms; **C** = company-level observations; **S-Y** = state-year observations; **S** = state-level observations; **HCFA** = Health Care Financing Administration; **NAIC** = National Assoc. of Insurance Commissioners; **MLM** = Medical Liability Monitor; **GP** = general practitioner; **class 3** = St. Paul Company Class 3 physician (family physician, minor surgery, no obstetrics); **mix** = internists, general surgeons and OBGYN; **log** = log-transformed dependent variable; **rse** = robust standard errors; **AR** = auto-regressive model; **CL** = clustered standard errors; **chk** = authors checked; **BS** = bootstrapped standard errors; **Q** = quantile regression; **n/a** = control not applicable. BLACK BOXES indicate methodologically sound choice. RED BOXES (OR GRAY if in black and white) indicate choice that might generate a biased estimate or other problems. WHITE BOXES indicate method/control not used.

Table 5c
Any cap; Loss ratios

[illegible]

Table 6a
Non-economic damages cap; Premiums paid by doctor

[illegible]

Table 7
Punitive damages cap

[illegible]

In studies that use the ratio of earned premiums to incurred losses as a proxy for premiums, disaggregating quantity sold from price is possible but controlling for the change in policy types sold is not. Given that tort reform likely impacts physician choices over coverage levels,³⁷ if regression results suggest that tort reform reduces the price per indemnified dollar, we are unable to discern whether amounts physicians paid for insurance decreased or whether the distribution of policy types purchased changed. In short, changes in premium-to-loss ratios imply nothing about changes in the absolute amount providers pay for insurance. Ratios are a better measure of profitability.

Finally, some researchers employ data on actual prices paid by physicians in exchange for coverage of a particular type (e.g., policies with limits of \$100,000 per occurrence and \$300,000 annually for general practitioners). Using this type of data allows us to separate quantity and price and to draw inferences about amounts physicians pay for insurance, but it does not allow for general conclusions about prices. For example, Zuckerman et al. (1990) employ data only on policies with \$100,000/\$300,000 limits for three different specialties—general practice, general surgery and obstetrics/gynecology. They report finding that “imposing a cap on physician liability—which in many cases is a limit on plaintiff recovery—lower premiums substantially” (Zuckerman et al., 1990, p. 180). Given the nature of the data, however, the authors (and all authors who use data on prices paid for particular types of policies) can claim only that caps are associated with lower premiums related to policies with those particular coverage levels.³⁸

³⁷ For example, a physician who believes that tort reform reduces liability exposure might be willing to switch from a policy that covers \$1M per occurrence to one that covers only \$500,000 per occurrence.

³⁸ Importantly, the authors over-generalize in a second way. The regression results suggest that caps on physician liability that allow for patients to collect amounts over the cap from patient compensation funds reduced premiums for \$100,000/\$300,000 policies for three specialties in the range of 13.6% to 16.9% but also that premiums were not impacted by caps on non-economic damages. This distinction is not mentioned in the conclusion section. General conclusions in discussion sections and abstracts often need to be taken with a grain of salt.

This would not be a problem if all policy types were impacted similarly by caps, but they likely are not. For example, prices for policies with relatively high limits are more likely to be impacted by caps because caps are more likely to bind amounts indemnified for claims brought against high-limit policies. Furthermore, if caps impact physician choices over policy types, changes in average risk levels and changes in demand across the mix of policy types will impact prices separate from the direct impacts of caps. Unless prices for a broader set of policy types are analyzed or these sorts of factors are controlled, generalizing from reported estimates is impossible. Moreover, without knowing more about the percentage of physicians who purchase the types of policies studied, getting a sense of the impact of caps on the market as a whole is impossible.³⁹

In addition, the Medical Liability Monitor dataset, the most common dataset used to estimate the impacts of caps on premiums paid for individual policies, comes with severe limitations (Mello, 2006). First, the majority of price information in the dataset relates to policies with \$1M/\$3M coverage for limited number of physician specialties. While we have very little information about the distribution of policies purchased by physicians, recently reported evidence suggests that a large segment of the physician population might purchase policies with less coverage (Zeiler et al., 2007). Second, the MLM collects the data by surveying insurers who voluntarily participate. Insurers who are less stable or have higher prices might be less likely to report to the MLM for fear of sending negative signals to consumers. Thus, the survey might generate an unrepresentative sample. Third, while the MLM reports premiums by county for some large states, most reports are average premiums of major insurers across counties. Given that premiums vary considerably by county, at least in some states, reporting averages might introduce substantial noise into the estimates.

³⁹ Mello (2006) notes similar shortcomings.

In sum, given the nature of the premiums data employed, every original study comes with severe limitations. No study is able to provide general estimates of the impact of caps on insurance prices.

2. Possible Model Assumption Violations Not Reliably Addressed

One of the most glaring problems with this literature is the failure to reliably address potential model assumption violations. The red boxes in the figures flag the estimates with specific problems. The Tables reveal that none of the 197 estimates was generated using methods that account for all the relevant basic assumptions. For example, none of the models employ controls for potential simultaneity even though it is theoretically plausible (and highly likely) that at least some state legislators implemented caps in response to premium trends. In addition, very few studies report even testing for assumption violations and fewer employ solid techniques to address observed or assumed violations (e.g., non-normal errors, heteroskedastic error terms and serial correlation). When violations are assumed but do not exist, employing techniques to address these non-existent violations can lead to unreliable estimates. In some cases techniques to address violations are employed but no evidence is provided to demonstrate that the violation was remedied (e.g., in some cases log-transforming the dependent variable will result in normally distributed errors, but it might not—the distribution must be examined).

3. Measurement Issues Abound

The Tables also reveal that measurement issues are prevalent. First, many studies do not account for effective dates, legislative repeals and judicial overturns. Many do not account for the possibility of insurers' delayed response to caps, and none accounts for possible anticipation effects. Almost all estimates are derived from models that do not properly weight observations to

account for market share and market size differentials, etc.⁴⁰ In many cases the potential for bias caused by outliers is either ignored or not adequately accounted for. Measurement errors can lead both to inaccurate estimates and incorrect conclusions regarding whether caps have a statistically significant impact on premiums.

4. Lack of Controls for Alternative Explanations

The Tables also signal vast potential for biased estimates due to the lack of control for alternative explanations. Given that reforms are regularly passed in packages, omitting controls for other tort reforms that put downward pressure on premiums will lead to false positives when it comes to assessing whether caps reduce premium growth. None of the studies that find a statistically significant negative impact on premiums controls for all (or indeed most) variables theorized to be both correlated with caps and drivers of premium variation. While some studies employ fixed effects, these control only for variables that vary over time and are fixed across cross sections (e.g., companies and states) and for variables that vary across cross sections and are constant across time. In some cases, fixed effects are not employed, which invites bias if fixed differences across states or fixed differences across time impact premiums.⁴¹

While several of the original studies employ a difference-in-differences model that controls for both time fixed effects and group fixed effects, none perform an analysis to determine whether states that did not implement caps during or prior to the study period make good control groups (Donohue and Ho, 2007). In addition, none check for pre-treatment trends, which would necessitate the use of a different empirical model.

⁴⁰ “Computing a simple average, rather than a weighted average, treats the companies as though they have identical shares of the market. Adjusting for market share and also for the number of physicians insured in each region of the state can make a big difference in the estimate of statewide average premiums: for Kentucky in 2002, for example, the simple average premium for obstetrician-gynecologists was \$58,287 but the weighted average premium was 19 percent less (\$48,897).” Mello (2006)

⁴¹ The political leaning of states is an example of a possible state fixed effect. Time fixed effects include federal interventions such as National Practitioner Databank reporting requirements, which some believe lead hospitals to encourage patients to drop physicians from claims so physicians can avoid reporting to the Databank

4. Conclusion

Tables 5-7 clearly demonstrate wide variation in methods employed to estimate the impacts of caps on premium growth. Variation in results likely is due, at least in part, to variation in data and method employed. While a more systematic technique referred to as meta-analysis exists for separating the effects of modeling techniques on results, we were not able to employ it here due to the small number of estimates relative to the number of methods we identified for our analysis. Thus, we are unable to make empirical claims about systematic impacts of different methods on estimates. The Tables, however, make a clear case for the use of an abundance of caution when deciding whether to put any amount of weight on any one study.

We detected no differences in methodological quality between results that support the caps-reduce-premiums claim and those that do not. One question we might ask is: which estimate is most reliable from a methods perspective? More generally, we might wish to get a sense of how much weight can be assigned to each estimate given the methods employed. Unfortunately, this question cannot be answered using the information we report. Even if some of the estimates were produced using nearly perfect methods, just one model violation or omitted variable might lead to more severe problems than those produced by models with several deficiencies.

In closing, we turn to the big picture. Most agree that our primary goals are to increase access to health care, improve its quality and to lower its cost (or to at least achieve efficient use of resources allocated to health care markets). These goals might suggest that we need to turn our attention to the impacts of damages caps on other important outcomes, including physician supply, defensive medicine, health care costs and health care insurance premiums. While these outcomes are important and studying them is necessary for a comprehensive view of the impact

of caps, proponents of caps continue to make claims about their impact on medical malpractice insurance premiums, and these claims continue to impact legislative outcomes and court decisions on the constitutionality of caps. It remains important, therefore, to get a solid handle on general inferences we can draw from the empirical literature that focuses on this albeit narrow question. Our analysis reveals not only that the results are mixed, despite claims to the contrary, but also that more methodologically sound research that employs more comprehensive data is essential.

References

- Angrist, Joshua D. and Jorn-Steffen Pischke (2009), *MOSTLY HARMLESS ECONOMETRICS: AN EMPIRICIST'S COMPANION*, (Princeton, NJ: Princeton University Press).
- Arlen, Jennifer and W. Bentley MacLeod (2005), *Torts, Expertise, and Authority: Liability of Physicians and Managed Care Organizations*, *RAND JOURNAL OF ECONOMICS*, 36, 494-519.
- Ashenfelter, Orley (1978), *Estimating the Effect of Training Programs on Earnings*, *REVIEW OF ECONOMICS AND STATISTICS*, 60, 47-57.
- Avraham, Ronen (Sept. 2011), *Database of State Tort Law Reforms (DSTLR 4th)*, University of Texas Law, Law and Econ Research Paper No. 184. Available at SSRN: <http://ssrn.com/abstract=902711> or <http://dx.doi.org/10.2139/ssrn.902711>.
- Avraham, Ronen and Max Schanzenbach (2010), *The Impact of Tort Reform on Private Health Insurance Coverage*, *AMERICAN LAW AND ECONOMICS REVIEW*, 12, 319-55.
- Barker, Drucilla K. (1992), *The Effects of Tort Reform on Medical Malpractice Insurance Markets: An Empirical Analysis*, *JOURNAL OF HEALTH POLITICS, POLICY AND LAW*, 17, 143-61.
- Bertrand, Marianne, Esther Duflo, & Sendhil Mullainathan (2004), *How Much Should We Trust Differences-in-Differences Estimates?*, *QUARTERLY JOURNAL OF ECONOMICS* 119, 249-75.
- Bhat Vasanthakumar N. (2001), *MEDICAL MALPRACTICE: A COMPREHENSIVE ANALYSIS*, (Westport: Auburn House).
- Blackmon, Glenn and Richard J. Zeckhauser (1991), *State Tort Reform Legislation: Assessing Our Control of Risks*, in Peter Shuck, ed., *TORT LAW AND THE PUBLIC INTEREST: COMPETITION, INNOVATION, AND CONSUMER WELFARE*, (New York: W.W. Norton and Company and the American Assembly).
- Born, Patricia H. and W. Kip Viscusi (1998), *The Distribution of the Insurance Market Effects of Tort Liability Reforms*, *BROOKINGS PAPERS ON ECONOMIC ACTIVITY: MICROECONOMICS*.
- Crain, Nicole V., W. Mark Crain, Lawrence J. McQuillan and Hovannes Abramyan (2009), *TORT LAW TALLY: HOW STATE REFORMS AFFECT TORT LOSSES AND TORT INSURANCE PREMIUMS*, (San Francisco: Pacific Research Institute).
- Danzon, Patricia M., Andrew J. Epstein and Scott J. Johnson (2004), *The 'Crisis' in Medical Malpractice Insurance*, In Richard Herring and Robert E. Litan, eds. *BROOKINGS-WHARTON PAPERS ON FINANCIAL SERVICES*, (Washington DC: Brookings Institution Press).
- Donohue, John J. III and Daniel E. Ho (2007), *The Impact of Damage Caps on Malpractice Claims: Randomization Inference with Difference-in-Differences*. *JOURNAL OF EMPIRICAL LEGAL STUDIES*, 4, 69-102.

- Gius, Mark P. (1998), *Using Panel Data to Estimate the Determinants of Medical Malpractice Premiums*, APPLIED ECONOMICS LETTERS, 5, 37-9.
- Grace, Martin F. and Tyler J. Leverty (2011), *How Tort Reform Affects Insurance Markets*, JOURNAL OF LAW, ECONOMICS AND ORGANIZATIONS (forthcoming).
- Greene, William H. (2002), *ECONOMETRIC ANALYSIS* (5th ed.) (New Jersey: Prentice Hall).
- Ho, Daniel E. (2005), *Why Affirmative Action Does Not Cause Black Students to Fail the Bar*, YALE LAW JOURNAL, 114, 1997-2004.
- Hsiao, C. (2003), *ANALYSIS OF PANEL DATA*. 2d ed. (Cambridge, U.K.: Cambridge University Press).
- Huber, Peter J. (1967), *The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions*, PROCEEDINGS OF THE FIFTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, vol. 1 (Berkeley CA: University of California Press).
- Hyman, David A., Bernard S. Black, Kathryn Zeiler, Charles M. Silver, and William M. Sage (2007), *Do Defendants Pay What Juries Award? Post-Verdict Haircuts in Texas Medical Malpractice Cases, 1988-2003*, JOURNAL OF EMPIRICAL LEGAL STUDIES, 4, 3-68.
- Hyman, David and William Sage (2011), *DO HEALTH REFORM AND MALPRACTICE REFORM FIT TOGETHER?* (Washington, DC: American Enterprise Institute for Public Policy Research).
- Kachalia, Allen and Michelle Mello (2011). *New Directions in Medical Liability Reform*, NEW ENGLAND JOURNAL OF MEDICINE, 364, 1564-72.
- Kane, Carol K. and David W. Emmons (2005), *THE IMPACT OF CAPS ON DAMAGES: HOW ARE MARKETS FOR MEDICAL LIABILITY INSURANCE AND MEDICAL SERVICE AFFECTED?*, (Chicago: American Medical Association).
- Kane, Carol K. and David W. Emmons (2007), *THE IMPACT OF LIABILITY PRESSURE AND CAPS ON DAMAGES ON THE HEALTHCARE MARKET: AN UPDATE OF RECENT LITERATURE*, (Chicago: American Medical Association).
- Kessler, Daniel P. (2006), *The Effects of the U.S. Malpractice System: A Review of the Empirical Evidence*, Unpublished manuscript.
- Kilgore, Meredith L., Michael A. Morrissey and Leonard J. Nelson (2006), *Tort Law and Medical Malpractice Premiums*, INQUIRY, 43, 255-70.
- King, Gary and Margaret Roberts (2012), *How Robust Standard Errors Expose Methodological Problems They Do Not Fix*, Unpublished Manuscript.

- Klick, Jonathan and Thomas Stratmann (2007), *Medical Malpractice Reform and Physicians in High-Risk Specialties*, JOURNAL OF LEGAL STUDIES, 36, S121-42.
- Malani, Anup and Reif, Julian (2011), *Accounting for Anticipation Effects: An Application to Medical Malpractice Tort Reform*, University of Chicago Law & Economics, Olin Working Paper No. 578, available at SSRN: <http://ssrn.com/abstract=1945419>.
- Mello, Michelle M. (2006), MEDICAL MALPRACTICE: IMPACT OF THE CRISIS AND EFFECT OF STATE TORT REFORMS (Princeton, NJ: Robert Wood Johnson Foundation).
- Mello, Michelle M. and Allen Kachalia. (2010) EVALUATION OF OPTIONS FOR MEDICAL MALPRACTICE: A REPORT TO THE MEDICARE PAYMENT ADVISORY COMMISSION, (Washington, D.C.: MedPAC).
- Michelle M. Mello, Amitabh Chandra, Atul A. Gawande and David M. Studdert (2010), *National Costs of The Medical Liability System*, HEALTH AFFAIRS, 29, 1569-77.
- Nelson, Leonard J., Michael A. Morrissey and Meredith Kilgore (2007), *Damage Caps in Medical Malpractice Cases*, THE MILBANK MEMORIAL QUARTERLY, 85, 259-86.
- RAND (2011), *Analysis of Medical Malpractice*, available at: http://www.rand.org/pubs/technical_reports/TR562z17/analysis-of-medical-malpractice.html.
- Rosenbaum, Paul and Rubin, Donald B. (1983), *The central role of the propensity score in observational studies for causal effects*, BIOMETRIKA 70 (1983): 41–55.
- Rosenbaum, Paul R (1984), *The Consequences of Adjusting for a Concomitant Variable that Has Been Affected by the Treatment*, JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES A, 147, 656-66.
- Rubin, Paul H. (1993), TORT REFORM BY CONTRACT (Washington, DC: American Enterprise Institute Press).
- Rubin, Paul H. and Joanna M. Shepherd (2008), *The Demographics of Tort Reform*, REVIEW OF LAW AND ECONOMICS, 4, 591-620.
- Sharkey, Catherine M. (2005), *Unintended Consequences of Medical Malpractice Damages Caps*, NEW YORK UNIVERSITY LAW REVIEW, 80, 391-512.
- Sloan, Frank A. (1985), *State Responses to the Malpractice Insurance 'Crisis' of the 1970s: An Empirical Assessment*, JOURNAL OF HEALTH POLITICS, POLICY AND LAW, 9, 629-46.
- Sloan, Frank A. and John H. Shadle (2009) *Is There Empirical Evidence for Defensive Medicine? A Reassessment*, JOURNAL OF HEALTH ECONOMICS, 28, 481-91.

- Studdert, David M. and Michelle M. Mello (2004), *Medical Malpractice*, NEW ENGLAND JOURNAL OF MEDICINE, 350, 283-92.
- Thomas, J. William, Erika C. Ziller and Deborah A. Thayer (2010), *Low Costs of Defensive Medicine, Small Savings from Tort Reform*, HEALTH AFFAIRS, 29, 1578-84.
- Thorpe, Kenneth E. (2004), *The Medical Malpractice Crisis: Recent Trends and the Impact of State Tort Reforms*, HEALTH AFFAIRS, 4, 20-30.
- U.S. Congress, Congressional Budget Office (2011), COST ESTIMATE HR5, HELP EFFICIENT, ACCESSIBLE, LOW-COST, TIMELY HEALTHCARE (HEALTH) ACT OF 2003 (Washington, DC: U.S. Government Printing Office).
- U.S. Congress, Congressional Budget Office (2004), THE EFFECTS OF TORT REFORM: EVIDENCE FROM THE STATES, (Washington, DC: U.S. Government Printing Office).
- U.S. Congress, Office of the Technology Assessment (1993), IMPACT OF LEGAL REFORMS ON MEDICAL MALPRACTICE COSTS, (Washington, DC: U.S. Government Printing Office).
- U.S. Department of Health and Human Services. Office of the Assistant Secretary for Planning and Evaluation (2002), CONFRONTING THE NEW HEALTH CARE CRISIS: IMPROVING HEALTH CARE QUALITY AND LOWERING COSTS BY FIXING OUR MEDICAL LIABILITY SYSTEM, (Washington, DC: U.S. Government Printing Office).
- U.S. General Accounting Office (2003), MEDICAL MALPRACTICE IMPLICATIONS OF RISING PREMIUMS ON ACCESS TO HEALTH CARE, (Washington, DC: U.S. Government Printing Office).
- Viscusi, Kip W., Richard J. Zeckhauser and Patricia H. Born (1993), *The Effect of the 1980s Tort Reform Legislation on General Liability and Medical Malpractice Insurance*, JOURNAL OF RISK AND UNCERTAINTY, 6, 165-86.
- Viscusi Kip W. and Patricia Born (1995), *Medical Malpractice Insurance in the Wake of Liability Reform*, JOURNAL OF LEGAL STUDIES, 24, 463-90.
- Viscusi, Kip W. and Patricia H. Born (2005), *Damage Caps, Insurability, and the Performance of Medical Malpractice Insurance*, JOURNAL OF RISK AND INSURANCE, 72, 23-43.
- White, Halbert (1980), *A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity*, ECONOMETRICA, 48, 817-30.
- Wooldridge, Jeffery M. (2008), INTRODUCTORY ECONOMETRICS: A MODERN APPROACH (4th ed.). (South-Western College Pub).

Yang, Y. Tony, David M. Studdert, S.V. Subramanian and Michelle M. Mello (2008), *A Longitudinal Analysis of the Impact of Liability Pressure on the Supply of Obstetrician-Gynecologists*, JOURNAL OF EMPIRICAL LEGAL STUDIES, 5, 21-53.

Zeiler, Kathryn (2003), *Medical Malpractice and Contract Disclosure: A Study of the Effects of Legal Rules on Behavior in Health Care Markets*. California Institute of Technology Dissertation.

Zeiler, Kathryn, Charles Silver, Bernard Black, David A. Hyman and William M. Sage (2007), *Physicians' Insurance Limits and Malpractice Payments: Evidence from Texas Closed Claims, 1990-2003*, JOURNAL OF LEGAL STUDIES, 36, S9-45.

Zuckerman, S., Randall R. Bovbjerg and Frank A. Sloan (1990), *Effects of Tort Reforms and Other Factors on Medical Malpractice Insurance Premiums*, INQUIRY, 27, 167-82.

Appendix 1: Scope of literature reviews

Review	Time period covered⁴²	Method for inclusion	Independent variables addressed	Dependent variables addressed	Number of studies included⁴³
OTA 1993	1990-1992	Studies employing systematic empirical methods, excluded single-state studies, studies pre-dating 1980, descriptive studies and literature reviews	Various tort reforms	Premiums, severity, frequency	3
CBO 2004	1993-2004	No explicit method	Various tort reforms	Premiums, brief mention of frequency and severity, defensive medicine	3
Kessler 2006	1990-1992	No explicit method	Various tort reforms (and other reforms)	Brief mention of premiums, claim severity, defensive medicine, various other aspects of the malpractice system	2
Kane 2005, 2007	1985-2006	Papers that employ statistical techniques to account for competing explanations	Focused on caps, brief mention of other reforms	Premiums, brief mention of frequency and severity, physician supply, defensive medicine	5
Nelson 2007	1990-2006	Better studies (control for differences across states and time and nonrandom enactment of the law). Strongest studies use fixed-effects.	Damages caps	Premiums, defensive medicine, physician supply	9
Mello 2004-2010	1985-2006	Specifically identified criteria for rigor, but included weak studies	Various tort reforms (and other reforms)	Premiums, frequency, severity, physician supply, defensive medicine, quality of care, defense costs	8
RAND 2011	1985-2006	Used Mello's selected studies	Various tort reforms	Premiums, frequency, severity, health expenditures	7

⁴² The time period covered includes only studies addressing the effect of damage caps on premiums. The time period reflects the range of publication dates covered by the primary studies included in the review.

⁴³ This number includes only studies addressing observed (as opposed to simulated) effects of medical malpractice damages caps on premiums. We did not count studies that group caps with other types of tort reforms even when the review author included these types of studies in the analysis. We also did not include primary studies that did not produce regression results.

Appendix 2: Studies of the impact of caps on premiums included in each literature review

	Sloan 1985	Zucker- man 1990	Black- mon 1991	Barker 1992	Viscusi 1993	Viscusi 1995	Born 1998	Gius 1998	Danzon 2004	Thorpe 2004	Viscusi 2005	Kil- gore 2006
OTA 1993		X	X	X								
CBO 2004					X		X			X		
Kes- sler 2006		X		X								
Nelson 2007		X	X ⁴⁴		X	X		X	X	X	X	X
Kane 2005, 2007	X	X								X	X	X
Mello 2004- 2010	X	X	X		X				X	X	X	X
RAND 2011	X	X	X		X				X	X	X	

Appendix 3: Search strategy

Our goal was to locate all relevant reported and unreported regression results included in studies of whether damages caps (of any sort) impact premiums. Table A3 summarizes the specific search strategies by database. When these search terms generated roughly 250 or fewer results, we evaluated each for relevance. We used full text search where available (and where it did not generate too many results) or, alternatively, we searched using keywords or subject headings. For several databases, we were able to use a very general search term “damage cap” OR “damages cap” OR “tort reform” without generating an unduly number of results. For two additional databases, the general phrase “tort reform” generated too many results, so we narrowed the search to: “damage cap” OR “damages cap” OR (“tort reform” AND (empirical OR regression)).

For several additional databases, the term “tort reform” (even when paired with these additional search terms) generated large numbers of results, as did “damage cap” OR “damages

⁴⁴ Nelson cites to a 1990 report, which was a less detailed study than the 1991 results reported in Blackmon’s book.

cap” without additional search terms. The search term (“damage cap” OR “damages cap”) AND (empirical OR regression) was used for several databases. For several additional databases, the term “empirical” generated too many results when paired with “damage cap” OR “damages cap,” we used the search term ((“damage cap” OR “damages cap”) AND regression).

Table A3: Search terms by database

Search terms	Databases
“damage cap” OR “damages cap” OR “tort reform”	SSRN, Econlit, Catalogue of US Government Publications, ProQuest Congressional
“damage cap” OR “damages cap” OR (“tort reform” AND (empirical OR regression))	PubMed, Index to Legal Periodicals
(“damage cap” OR “damages cap”) AND (empirical OR regression)	ProQuest Dissertations and Theses, JSTOR, Academic Search Primer, ABI Inform Complete, BePress
(“damage cap” OR “damages cap”) AND regression	Hein Online, Westlaw (law review database), Google Books and Google Scholar*
(“damage cap” OR “damages cap”) AND regression AND “medical malpractice”)	Google

* Additional searches performed in Google Scholar were performed using the following terms: noneconomic damages caps medical malpractice insurance regression, tort reform regression damage caps medical malpractice, claim severity tort reform regression, claim frequency tort reform regression, health insurance premiums tort reform regression, tort reform regression, liability insurance regression medical malpractice.

Finally, we used the most specific search term ((“damage cap” OR “damages cap”) AND regression AND “medical malpractice”) in Google. This generated nearly 700 results, and we reviewed the first 250. We also performed a number of Google Scholar searches using the following search terms individually: damages caps, medical malpractice insurance regression, tort reform regression damage caps medical malpractice, claim severity tort reform regression, claim frequency tort reform regression, health insurance premiums tort reform regression, tort reform regression, liability insurance regression medical malpractice.

An additional search was performed using the following databases: NBER, Wiley Online Library, Science Direct, Lexis Nexis, Proquest, Health Affairs, and the Georgetown Law Library’s E-Journal Finder). Search terms employed included: damage caps, medical malpractice damage caps, empirical AND tort AND malpractice, insurance AND tort liability, empirical

AND tort reform, medical malpractice, medical malpractice AND tort reform, damage caps
AND malpractice AND reform, caps AND damages, and summary((damage /3 cap) OR tort
reform /100 economic! and empirical). To supplement this search, we reviewed the
bibliographies of the primary studies and the literature reviews.